

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 07-334507
 (43)Date of publication of application : 22.12.1995

(51)Int.Cl. G06F 17/28
 G06T 13/00
 G10L 3/00

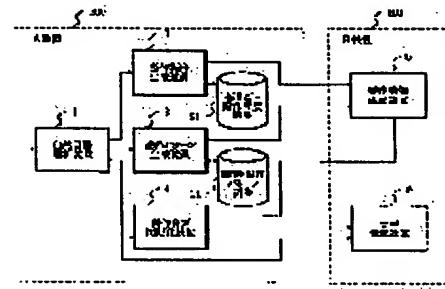
(21)Application number : 06-126002 (71)Applicant : NEC CORP
 (22)Date of filing : 08.06.1994 (72)Inventor : RO SAN

(54) HUMAN BODY ACTION AND VOICE GENERATION SYSTEM FROM TEXT

(57)Abstract:

PURPOSE: To input a text written in a natural language, to generate a human body action which is matched with the content of the text and to output the content of the text by means of voice.

CONSTITUTION: A natural language analysis device 1 fetching a word from the text, a verb/action pattern dictionary 21 describing the corresponding relation of the word showing the action and a human body action pattern, an action pattern generation device 2 retrieving the verb/action pattern dictionary 21 and generating the human body pattern, a modifier/action degree dictionary 31 describing the corresponding relation of a modifier and the degree of the action, an action degree generation device 3 retrieving the modifier/action degree dictionary 31 and generating the action degree of the human body pattern, an action and voice synchronization device 4 outputting operation time data obtained by making the appearing position of a verb in the text corresponding to the start time of action generation, and outputting voice data, an action video generation device 5 generating time sequential video data of the human body action and video-outputting it and a voice synthesis device 6 outputting voice by a regular voice synthesis method are provided.



LEGAL STATUS

[Date of request for examination] 08.06.1994

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number] 2976811

[Date of registration] 10.09.1999

[Number of appeal against examiner's decision]

[of rejection]

[Date of requesting appeal against examiner's
decision of rejection]

[Date of extinction of right]

(19)日本国特許庁 (JP)

(12) 公開特許公報 (A)

(11)特許出願公開番号

特開平7-334507

(43)公開日 平成7年(1995)12月22日

(51)Int.Cl. ⁶	識別記号	序内整理番号	F I	技術表示箇所
G 06 F 17/28				
G 06 T 13/00				
G 10 L 3/00	S			
		8420-5L	G 06 F 15/ 38	
		0834-5H	15/ 62	3 4 0 A
			審査請求 有	請求項の数3 OL (全7頁)

(21)出願番号 特願平6-126002

(22)出願日 平成6年(1994)6月8日

(71)出願人 000004237

日本電気株式会社

東京都港区芝五丁目7番1号

(72)発明者 呂山

東京都港区芝五丁目7番1号 日本電気株式会社内

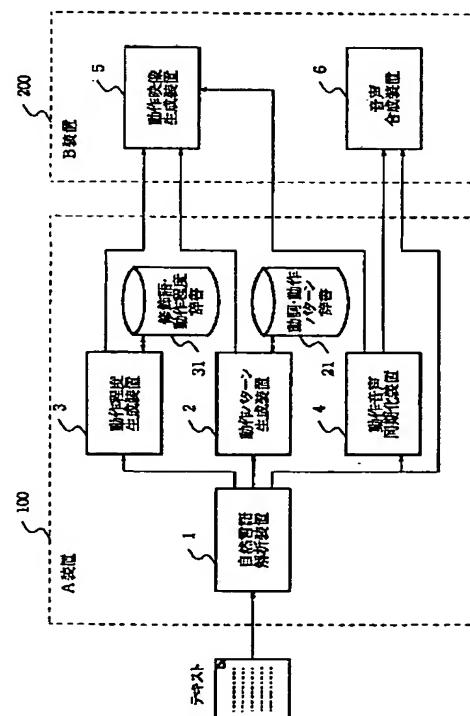
(74)代理人 弁理士 京本直樹 (外2名)

(54)【発明の名称】 テキストから的人体動作音声生成システム

(57)【要約】

【目的】 自然言語で書かれたテキストを入力し、テキストの内容に合致するような人体動作を生成するとともに、音声でテキストの内容を出力する。

【構成】 テキストを取り出す自然言語解析装置1と、動作を表す単語と人体動作パターンの対応関係を記述する動詞・動作パターン辞書21と、動詞・動作パターン辞書21を検索し人体動作パターンを生成する動作パターン生成装置2と、修飾語と動作の程度の対応関係を記述する修飾語・動作程度辞書31と、修飾語・動作程度辞書31を検索し人体動作パターンの動作程度を生成する動作程度生成装置3と、テキスト中の動詞の出現位置を動作生成の開始時刻に対応付けた動作時間データおよび音声時間データを出力する動作音声同期化装置4と、人体動作の時系列映像データを生成し映像出力する動作映像生成装置5と、規則音声合成方法で音声を出力する音声合成装置6とから構成される。



1

【特許請求の範囲】

【請求項1】 テキストから動詞や副詞などの単語を取り出す自然言語解析手段と、動詞などの動作を表す単語と人体動作パターンの対応関係を記述する動詞・動作パターン辞書と、前記自然言語解析手段で抽出された動詞を用いて前記動詞・動作パターン辞書を検索し人体動作パターンを生成する動作パターン生成手段と、動詞を修飾する副詞などの修飾語と動作の程度の対応関係を記述する修飾語・動作程度辞書と、前記自然言語解析手段で抽出された修飾語を用いて前記修飾語・動作程度辞書を検索し前記人体動作パターンの動作程度を生成する動作程度生成手段と、動作映像出力と合成音声出力とを同期させるためのテキスト中の動詞の出現位置を動作生成の開始時刻に対応付けた動作時間データ、およびテキストの長さから計算されるテキストの読み上げ時間を含む音声時間データを出力する動作音声同期化手段と、前記人体動作パターン、動作程度および動作時間データを含む動作生成命令を入力とし人体動作の時系列映像データを生成し表示手段に出力する動作映像生成手段と、前記テキストおよび前記音声時間データを含む音声生成命令を入力とし規則音声合成方法で音声を出力する音声合成手段とを備えることを特徴とするテキストから的人体動作音声生成システム。

【請求項2】 請求項1記載のテキストから的人体動作音声生成システムにおいて、前記動作パターン生成手段、前記動作程度生成手段および前記動作音声同期化手段から出力される前記動作生成命令と前記音声生成命令とを文章動作記述ファイルに変換する動作音声生成命令・テキスト変換手段と、前記動作音声生成命令・テキスト変換手段に変換された文章動作記述ファイルを格納する文章動作記述ファイル蓄積手段と、前記文章動作記述ファイル蓄積手段からの文章動作記述ファイルを前記動作生成命令と音声生成命令とに変換し前記動作映像生成手段および前記音声合成手段に出力するテキスト・動作音声生成命令変換手段とを備え、外部のエディタで前記文章動作記述ファイルを修正できるようにしたことを特徴とするテキストから的人体動作音声生成システム。

【請求項3】 前記文章動作記述ファイルは人間可読のテキストであることを特徴とする請求項2記載のテキストからの人体動作音声生成システム。

【発明の詳細な説明】**【0001】**

【産業上の利用分野】 本発明は、テキストからの人体動作音声生成システムに関し、特に計算機システムを用いて自然言語で書かれたテキスタイルから、人体動作および音声の生成や人間の動きと音声のアニメーション

2

の作成を行うシステムに関する。

【0002】

【従来の技術】 従来の技術として、計算機による自然言語の構文解析技術を利用した人体動作の生成方法が、特開平4-264972号公報に開示されている。これは自然言語の構文解析技術を用いて、テキストから解析される動作を表す特定な単語などを利用して動作プログラムを生成し、人体動作の生成を行うものである。

【0003】 また、別の従来の技術として、音声合成装置

からの音素を利用して人間の口形の変化を生成する方法が、特開平2-234285号公報に開示されている。これはテキスタイルから規則音声合成装置によって生成された音声の音素を利用して、各音素に対応する口形特徴のパラメータを制御することにより、人間の口形の変化を生成するものである。

【0004】

【発明が解決しようとする課題】 上述した従来の技術には、次のような問題点が存在する。

【0005】 (1) 前者の場合は、テキストから動作の

生成が可能であるが、その動作に同期した音声の出力はできない。

【0006】 (2) さらに、前者では動作を記述するアニメーション専用の動作プログラムを生成するが、その動作プログラムが計算機プログラムと同様な形式に記述されているため、プログラマーではない一般利用者の編集作業には適していない。

【0007】 (3) 後者の場合は、テキストの音声出力から人間動作の一部である口形の変化だけを生成しており、より自然な人間な動作映像を生成するための身体などの他の部分の動きの生成が困難である。

【0008】

【課題を解決するための手段】 本発明のテキストからの人体動作音声生成システムは、上述した(1)項および(3)項に記載の課題を解決するために、テキストから動詞や副詞などの単語を取り出す自然言語解析手段と、動詞などの動作を表す単語と人体動作パターンの対応関係を記述する動詞・動作パターン辞書と、前記自然言語解析手段で抽出された動詞を用いて前記動詞・動作パターン辞書を検索し人体動作パターンを生成する動作パターン生成手段と、動詞を修飾する副詞などの修飾語と動作の程度の対応関係を記述する修飾語・動作程度辞書と、前記自然言語解析手段で抽出された修飾語を用いて前記修飾語・動作程度辞書を検索し前記人体動作パターンの動作程度を生成する動作程度生成手段と、動作映像出力と合成音声出力とを同期させるためのテキスト中の動詞の出現位置を動作生成の開始時刻に対応付けた動作時間データ、およびテキストの長さから計算されるテキストの読み上げ時間を含む音声時間データを出力する動作音声同期化手段と、前記人体動作パターン、動作程度および動作時間データを含む動作生成命令を入力とし人

体動作の時系列映像データを生成し表示手段に出力する動作映像生成手段と、前記テキストおよび前記音声時間データを含む音声生成命令を入力とし規則音声合成方法で音声を出力する音声合成手段とを備えている。

【0009】まだ、上述した(2)項に記載の課題を解決するために、上記の構成に前記動作パターン生成手段、前記動作程度生成手段および前記動作音声同期化手段から出力される前記動作生成命令と前記音声生成命令とを人間可読のテキストである文章動作記述ファイル変換する動作音声生成命令・テキスト変換手段と、前記動作音声生成命令・テキスト変換手段に変換された文章動作記述ファイルを格納する文章動作記述ファイル蓄積手段と、前記文章動作記述ファイル蓄積手段からの文章動作記述ファイルを前記動作生成命令と音声生成命令とに変換し前記動作映像生成手段および前記音声合成手段に出力するテキスト・動作音声生成命令変換手段とを付加することにより、外部のエディターで前記文章動作記述ファイルを修正できることを可能にしている。

【0010】

【作用】本発明においては、入力されたテキストを解析して語句ごとに分割し、動詞や動詞修飾語などの単語を取り出す。そして、動詞に出現する位置を動作開始の信号とする規則に基づいて、動作生成のタイミングを決定する。

【0011】また、動詞の種類に対応する人体動作パターンを決定したり、修飾語などを用いて、動作の動きの程度を決めるところがポイントである。これにより、テキストを与えると音声出力および音声と同期した円滑な人間の動作を自動的に作成することができる。

【0012】

【実施例】次に、本発明について図面を参照して説明する。

【0013】図1は、本発明のテキストから的人体動作音声生成システムの一実施例を示すブロック図であって、第1の発明の実施例の構成を示す図である。

【0014】本実施例のテキストからの人体動作音声生成システムは、図1に示すように、A装置100とB装置200とから成り、A装置100は、テキストから動詞や副詞などの単語を取り出す自然言語解析装置1と、動詞などの動作を表す単語と人体動作パターンの対応関係を記述する動詞・動作パターン辞書21と、自然言語解析装置1で抽出された動詞を用いて動詞・動作パターン辞書21を検索し人体動作パターンを生成する動作パターン生成装置2と、動詞を修飾する副詞などの修飾語と動作の程度の対応関係を記述する修飾語・動作程度辞書31と、自然言語解析装置1で抽出された修飾語を用いて修飾語・動作程度辞書31を検索し前記人体動作パターンの動作程度を生成する動作程度生成装置3と、動作映像出力と合成音声出力を同期させるためのテキスト中の動詞の出現位置を動作生成の開始時刻に対応付け

た動作時間データ、およびテキストの長さから計算されるテキストの読み上げ時間を含む音声時間データを出力する動作音声同期化装置4とから構成される。

【0015】また、B装置200は、人体動作パターン、動作程度および動作時間データを含む動作生成命令を入力とし人体動作の時系列映像データを生成し表示装置に出力する動作映像生成装置5と、テキストおよび音声時間データを含む音声生成命令を入力とし規則音声合成方法で音声を出力する音声合成装置6とから構成される。

【0016】自然言語解析装置1は、外部より入力されるテキストから個々の単語を取り出す。この自然言語解析装置1は、従来からの構文解析の手法を用いて、文から単語を抽出する。ここでは、CYK法を用いた構文解析の方法(杉村領一、赤坂宏二、久保幸弘：論理型形態素解析LAX, Proc. of the Logic Programming Conf., ICOT, 213-222, 1988年)などを利用している。なお、自然言語解析装置1は、既存技術を利用するので、ここでは詳細な説明を省略する。

【0017】この自然言語解析装置1によって、例えば、「彼が気持ち良く笑った」のような文を、「彼・が・気持ち良く・笑った」の個々の独立した語句に分解することができる。また、次に示すように、文の中に各々の語句の文法上の意味が得られる。

【0018】

彼 : 主語
が : 助詞
気持ち良く : 修飾語

笑った : 動詞

自然言語解析装置1から出力される動詞を動作パターン生成装置2の入力として、人体動作パターンを生成する。具体的には動詞および人体動作パターンの対応関係を格納した動詞・動作パターン辞書21を調べ、入力された動詞に対応した動作パターンを生成し出力する。

【0019】次に、人体動作パターンを生成する処理の詳細について説明する。

【0020】動詞・動作パターン辞書21は、1つの動詞に対して複数の人体動作パターンを対応させ、さらに、それぞれの人体動作パターンに優先度が付与されている。従って、動詞から人体動作パターンを検索するとき、優先度の高い順に人体動作パターンを出力する。

【0021】図5は、図1の動作パターン生成装置2の詳細を示すブロック図である。図5において、動作パターン制約条件辞書26は、順序を付けられた2つの前後の人体動作パターンが適切であるか否かの情報を格納する。例えば、頭部が左に傾けた状態で、頭部を前後に振るといった動作パターンは明らかに不自然で適切ではないと定義される。こうした前人体動作パターンから現在生成される人体動作パターンが適切であるか否かの知識

を動作パターン制約条件辞書26の中に定義する。

【0022】図6は、図5の動作パターン生成装置2が人体動作パターンを生成する処理の流れを示す流れ図である。

【0023】まず、自然言語解析装置1から得られた動詞を、動作パターン探索部22に入力する。このとき、カウント発生器23から現在のカウント値（初期値が“0”である）に“1”を加え、そのカウント値を動作パターン探索部22に送る。

【0024】以下、動作パターン生成装置2の動作について、カウント値=1の場合と、カウント値>1の場合とに分けて説明する。

①カウント値=1の場合；まず、動作パターン探索部22が入力された動詞を基に、動詞・動作パターン辞書21から優先度の最も高い人体動作パターンを取り出す。そして、取り出された人体動作パターンを出力すると同時に、この人体動作パターンとカウント値とを合わせてヒストリー記憶部24に記憶させる。

②カウント値>1の場合；動作パターン探索部22が入力された動詞を基に、動詞・動作パターン辞書21から優先度の最も高い人体動作パターン(MP_{i1})を取り出す。次に、ヒストリー記憶部24から現カウント値の1つ前の人體動作パターン(MP_{i-1})を取り出す。そして、人体動作パターン(MP_i)および(MP_{i-1})を動作パターン照合部25に送る。

【0025】動作パターン照合部25は、動作パターン制約条件辞書26を参考にし、1つ前の人體動作パターン(MP_{i-1})に対して現在の人體動作パターン(MP_i)が適切であるか否かを判断する。

【0026】そして、適切であると判断されると、現在の人體動作パターン(MP_i)を出力し、現在のカウント値と現在の人體動作パターン(MP_i)とをヒストリー記憶部24に記憶させる。また、適切ではないと判断されると、動作パターン探索部22がもう一度次に優先度の高い人体動作パターン(MP'_{i-1})を取り出す。

【0027】次に、この人體動作パターン(MP')と1つ前の人體動作パターン(MP_{i-1})とを用いて人體動作パターンの照合を行い、適切と判断されるまで人體動作パターンの探索・照合の処理を繰り返すとともに、探索されたすべての人體動作パターンが不適切であると判断されると、優先度の最も高い人體動作パターンを出力する。

【0028】さらに、自然言語解析装置1から出力される修飾語を動作程度生成装置3の入力として、人體動作パターンの程度を記述する動作程度のデータを生成する。具体的には、修飾語および人體動作パターンの動作程度の対応関係を格納した修飾語・動作程度辞書31を検索し、入力された修飾語に対応した動作程度を生成し出力する。ここで、動作程度を表現するため、数値データを用いることができる。

【0029】動作音声同期化装置4は、自然言語解析装置1から得られた語句を基に、動作の生成および合成音声の出力のタイミングを一致させる機能を有する。本発明においては、自然言語解析装置1からの句読点ではさまれた1区切りのテキストを動作生成と音声合成の基本単位として考え、このテキスト中に動作を表す動詞の出現する位置を動作生成の開始位置とする。その具体的な処理について以下に説明する。

【0030】図4は、動作音声同期化装置4の出力する動作・音声の時間データの一例を示す図である。図4において、まず、音声合成装置6の音声出力速度を基にして1語句を出力する総時間を計算する。例えば、音声合成装置6が1文字を出力するための所要時間をt秒とすると、

n 個の文字からなる語句の出力時間 $T_s = n \times t$ 秒となる。そして、語句の始めの単語に対応して、音声の開始時刻を t_{s0} 秒（相対時間が0秒である）とすると、音声の終了時刻 = $t_{s0} + T_s$ 秒となる。また、語句の中に動詞の出現する位置がm文字目にあるとすると、

動作生成の開始時間 $t_m = t_{s0} + m \times t$ 秒

動作時間の長さ $T_m = (n - m) \times t$ 秒

となる。図4には、このようにして計算された動作時間データと音声時間データとを示す。

【0031】次に、動作映像生成装置5は、動作パターン生成装置2からの人體動作パターン、動作程度生成装置3からの動作程度、および動作音声同期化装置4からの動作時間データを含む動作生成命令を入力として、ディスプレイ装置やVTRなどに人體動作の時系列画像を出力する。この動作映像生成装置5においては、人體動作パターンを複数の動作モジュールの合成による生成方式（例えば、呂山、吉坂主司、宮井均：「人體動作生成システムの提案座化」、情報処理学会第47回全国大会講演論文集（2），345-346，1993年）を利用する。

【0032】続いて、音声合成装置6については、既存の音声規則合成手法を利用することができる（山本誠一、樋口宜男、清水水徹：「テキスト編集機能付き音声規則合成装置の試作」、電子情報通信学会技術報告SP87-137, 1988年3月）。そして、自然言語解析装置1からの語句と動作音声同期化装置4からの時間データを含む音声出力命令を入力として、音声を合成し出力する。

【0033】図2は、第2の発明の一実施例の構成を示すブロック図である。本実施例は、図2に示すように、A装置100から出力される動作生成命令と音声生成命令とを人間可読のテキストである文章動作記述ファイル変換する動作音声生成命令・テキスト変換装置7と、この動作音声生成命令・テキスト変換装置7に変換された文章動作記述ファイルを格納する文章動作記述ファイル

蓄積装置 8 と、この文章動作記述ファイル蓄積装置 8 からの文章動作記述ファイルを動作生成命令と音声生成命令とに変換し、B 装置 200 に出力するテキスト・動作音声生成命令変換装置 9 とから構成される。

【0034】なお、A 装置 100 は、動作パターン生成装置 2、動作程度生成装置 3 および動作音声同期化装置 4 を含み、B 装置 200 は、動作映像生成装置 5 および音声合成装置 6 を含んでいるが、これら A 装置 100 および B 装置 200 については、第 1 の発明の実施例において既に説明済みであり、重複を避けるために省略し、図 2 の他の部分について説明する。

【0035】本実施例では、A 装置 100 から出力される動作生成命令と音声生成命令とを動作音声生成命令・テキスト変換装置 7 により文章動作記述ファイルのフォーマットに合ったテキストファイルに変換する。

【0036】図 3 は、このフォーマットの一例を示す図である。図 3において、テキストファイルに書き込まれたテキスト文章に対し、このテキスト中に含まれる動詞と同じ位置に、アンダーラインマークを付け、さらに、そのアンダーラインマークの下に A 装置 100 が生成した人体動作パターン名、人体動作パターンの動作程度のパラメータ p と動作時間の長さ t を記述する。

【0037】次に、動作音声生成命令・テキスト変換装置 7 により変換された文章動作記述ファイルを磁気ディスク装置などの外部記憶装置から構成される文章動作記述ファイル蓄積装置 8 に格納する。この格納した文章動作記述ファイルは、可読なテキストファイルの形式であるため、一般に市販されているテキストエディターを利用して、動作の修正などを容易に行うことができる。

【0038】テキスト・動作音声生成命令変換装置 9 は、動作音声生成命令・テキスト変換装置 7 とは反対に、文章動作記述ファイル書き込まれた人体動作パターン、人体動作パターンの動作程度および動作時間データを読み出し、B 装置 200 中の動作映像生成装置 5 に入力の動作生成命令に変換する。続いて、文章動作記述ファイル書き込まれたテキストを読み出し、動作音声同期化装置 4 で用いられた音声時間のデータの生成方法を利用して、音声出力用のテキストと音声時間データを生成し、音声合成装置 6 に入力し音声の出力を行う。

【0039】

【発明の効果】以上説明したように、本発明のテキストからの人体動作音声生成システムは、入力された自然言語のテキストを合成音声で出力し、音声と同期が取れた人体動作を自動的に生成することができる。

【0040】また、元のテキストに近い形の動作音声記述ファイルを作成し、普通のテキストエディターでそのファイルを編集することにより、最終的に生成される人体動作の調整を行うことができる。

【図面の簡単な説明】

10 【図 1】第 1 の発明の一実施例の構成を示すブロック図である。

【図 2】第 2 の発明の一実施例の構成を示すブロック図である。

【図 3】文章動作記述ファイルのフォーマットの一例を示す図である。

【図 4】本実施例の動作音声同期化装置の出力する動作・音声の時間データの一例を示す図である。

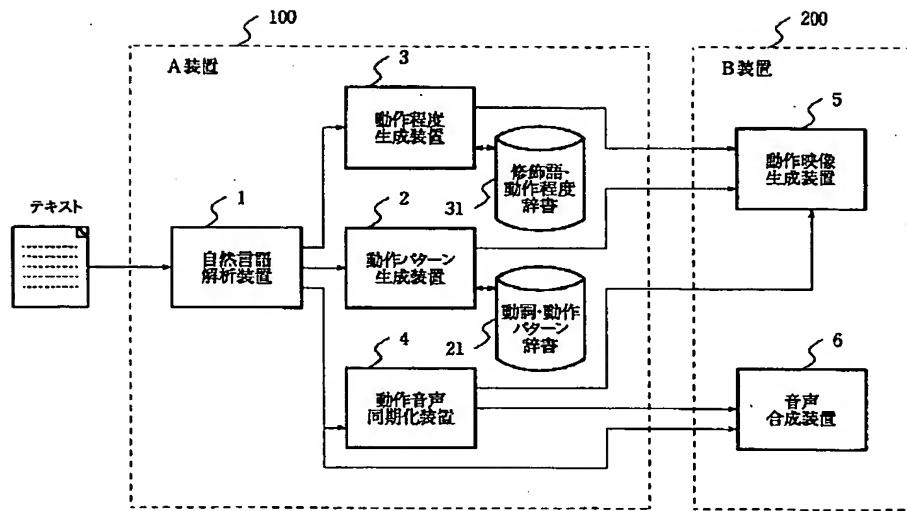
【図 5】図 1 の動作パターン生成装置の詳細を示すブロック図である。

20 【図 6】図 5 の動作パターン生成装置が人体動作パターンを生成する処理の流れを示す流れ図である。

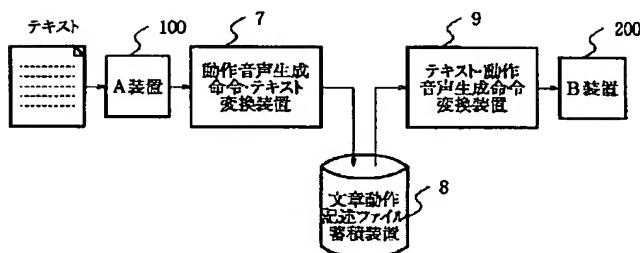
【符号の説明】

- | | |
|-----|-------------------|
| 1 | 自然言語解析装置 |
| 2 | 動作パターン生成装置 |
| 3 | 動作程度生成装置 |
| 4 | 動作音声同期化装置 |
| 5 | 動作映像生成装置 |
| 6 | 音声合成装置 |
| 7 | 動作音声生成命令・テキスト変換装置 |
| 30 | 文章動作記述ファイル蓄積装置 |
| 8 | テキスト・動作音声生成命令変換装置 |
| 9 | |
| 21 | 動詞・動作パターン辞書 |
| 22 | 動作パターン探索部 |
| 23 | カウント発生器 |
| 24 | ヒストリー記憶部 |
| 25 | 動作パターン照合部 |
| 26 | 動作パターン制約条件辞書 |
| 31 | 修飾語・動作程度辞書 |
| 100 | A 装置 |
| 40 | B 装置 |

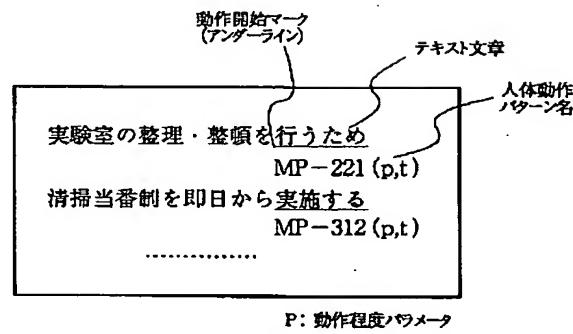
【図1】



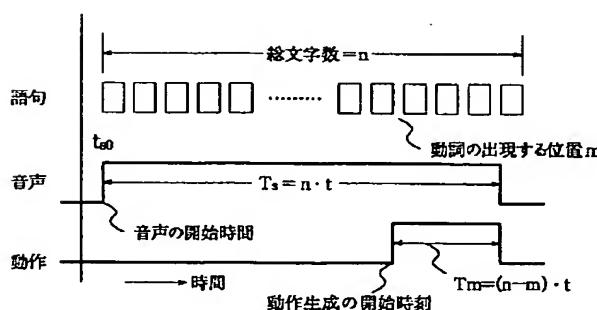
【図2】



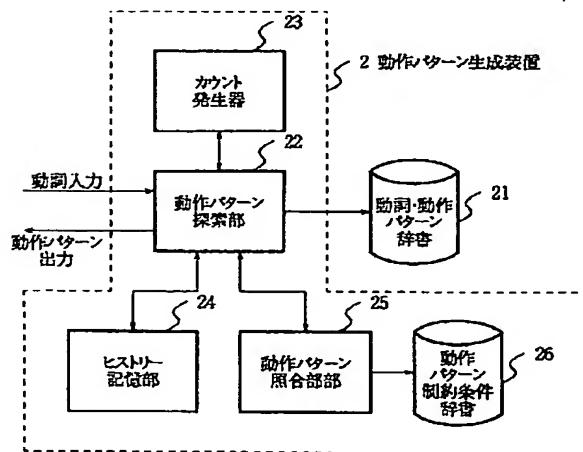
【図3】



【図4】



【図5】



【図6】

